Automated Metadata Population Service

(AMPS)

Spiral 1 Workshop

Mark Uhart, CKM , CSC

Verlynda Dobbs, Ph.D., Atlantic Consulting Services, Inc.

30 October
2008

AMPS Presentation Outline

- Background
- DoD Discovery Metadata Specification (DDMS)
- AMPS Implementation and Functionality
- Army Participation in Spiral 1 Testing
- AMPS Operational Example
- Documentation

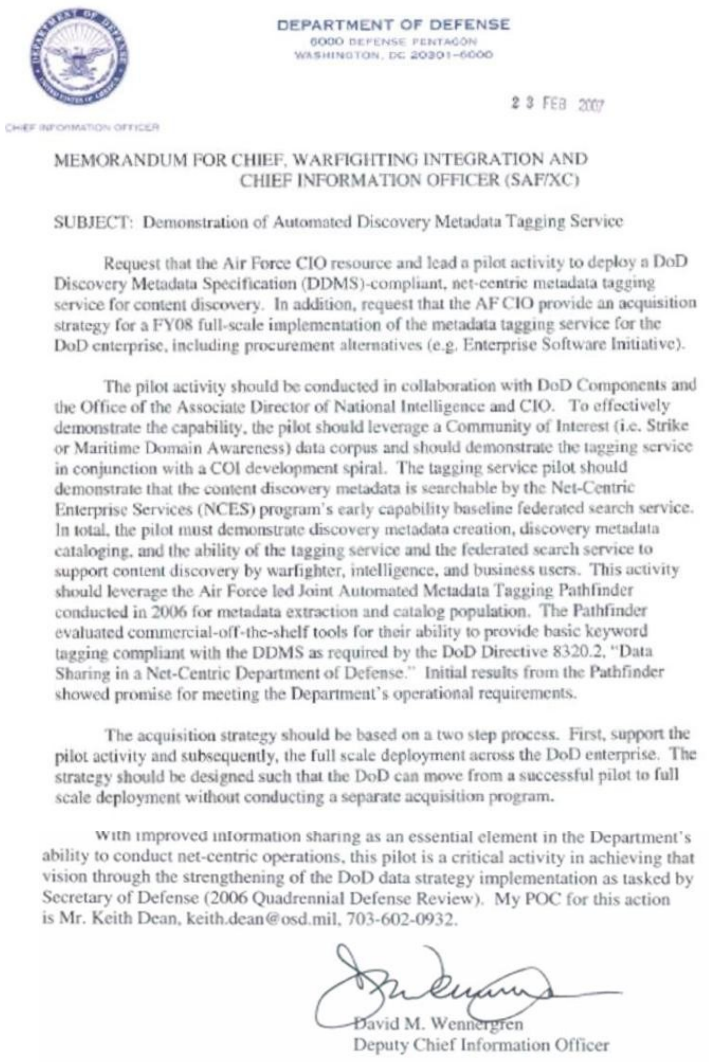
UNCLASSIFIED

Close Collaboration and K-Transfer

Automated Metadata Population Service (AMPS)

- **DoD Memorandum for Pilot activity – February 2007**
  - **Deploy DDMS compliant Service**
  - **Support metadata creation, metadata cataloging, and content discovery**
  - **Leverage the Pathfinder effort**
- **AMPS Working Group convened - April 2007**

Program Inception



DEPARTMENT OF DEFENSE
6000 DEFENSE PENTAGON
WASHINGTON, DC 20301-6000

CHIEF INFORMATION OFFICER

23 FEB 2007

MEMORANDUM FOR CHIEF, WARFIGHTING INTEGRATION AND
CHIEF INFORMATION OFFICER (SAF/XC)

SUBJECT: Demonstration of Automated Discovery Metadata Tagging Service

Request that the Air Force CIO resource and lead a pilot activity to deploy a DoD Discovery Metadata Specification (DDMS)-compliant, net-centric metadata tagging service for content discovery. In addition, request that the AF CIO provide an acquisition strategy for a FY08 full-scale implementation of the metadata tagging service for the DoD enterprise, including procurement alternatives (e.g. Enterprise Software Initiative).

The pilot activity should be conducted in collaboration with DoD Components and the Office of the Associate Director of National Intelligence and CIO. To effectively demonstrate the capability, the pilot should leverage a Community of Interest (i.e. Strike or Maritime Domain Awareness) data corpus and should demonstrate the tagging service in conjunction with a COI development spiral. The tagging service pilot should demonstrate that the content discovery metadata is searchable by the Net-Centric Enterprise Services (NCES) program's early capability baseline federated search service. In total, the pilot must demonstrate discovery metadata creation, discovery metadata cataloging, and the ability of the tagging service and the federated search service to support content discovery by warfighter, intelligence, and business users. This activity should leverage the Air Force led Joint Automated Metadata Tagging Pathfinder conducted in 2006 for metadata extraction and catalog population. The Pathfinder evaluated commercial-off-the-shelf tools for their ability to provide basic keyword tagging compliant with the DDMS as required by the DoD Directive 8320.2, "Data Sharing in a Net-Centric Department of Defense." Initial results from the Pathfinder showed promise for meeting the Department's operational requirements.

The acquisition strategy should be based on a two step process. First, support the pilot activity and subsequently, the full scale deployment across the DoD enterprise. The strategy should be designed such that the DoD can move from a successful pilot to full scale deployment without conducting a separate acquisition program.

With improved information sharing as an essential element in the Department's ability to conduct net-centric operations, this pilot is a critical activity in achieving that vision through the strengthening of the DoD data strategy implementation as tasked by Secretary of Defense (2006 Quadrennial Defense Review). My POC for this action is Mr. Keith Dean, keith.dean@osd.mil, 703-602-0932.

David M. Wennergren
Deputy Chief Information Officer

- **Air Force formed Automated Metadata Population Service (AMPS) Working Group**
- **NSA formed Information Assurance sub-group**
- **Participation**
  - **Government**
    - **Air Force**
    - **JFCOM**
    - **NSA**
    - **Army (BCKS)**
    - **DISA**
    - **Navy**
    - **DIA**
    - **NGA**
  - **Industry**
    - **Booz Allen Hamilton**
    - **eCompex**
    - **MITRE**
    - **Apache**

DoD Discovery Metadata Specification

- **The DoD Net-Centric Data Strategy (NCDS) and Directive 8320.2 require data sharing across the DoD, including the creation of new information resources to describe available data:**
- **[POLICY] 4.2. Data assets shall be made visible by creating and associating metadata (“tagging”), including discovery metadata, for each asset. Discovery metadata shall conform to the Department of Defense Discovery Metadata Specification (DDMS). [Department of Defense Directive Number 8320.2 (December 2, 2004), p. 2., directive certified current as of April 23, 2007]**
- **Use of DDMS is required!**
- **http://metadata.dod.mil/mdr/irs/DDMS/#DDMS_info**

Implementation - Goal

- **Provide a working instance of a metadata population framework to populate DDMS metacards for COIs**
- **Sufficiently flexible to allow incorporation of:**
  - **COI-specific business rules**
  - **Government-authored technologies**
  - **COTS technologies**

Implementation – Web Service

- **Possible to deploy:**
  - **in variety of environments (including laptop)**
  - **with restricted computing resources**
- **Exploits vocabulary products, specifically those that exhibit ontology characteristics such as class-subclass relations, synonymy and logical triples**

UNCLASSIFIED

Implementation – Open Source

- **Unstructured Information Management Architecture (UIMA) developed by IBM. An open source framework for analyzing asset contents and creating annotations. UIMA is in the process of becoming an OASIS normalized standard. [Apache Software Foundation. Apache UIMA, http://incubator.apache.org/uima**
- **Web Ontology Language (OWL)**
- **Web Service Description Language (WSDL)**
- **OpenOffice to process Microsoft Office files**

Functionality of AMPS

- **Inputs:**
  - **Data assets (Microsoft Office products, pdf, email, xml, etc)**
  - **Vocabularies (English dictionary, COI dictionaries, thesaurus)**
- **Outputs:**
  - **Metadata in DoD Discovery Metadata Specification (DDMS) format**
- **Mode of operation:**
  - **Content Manager User Interface - process one asset at a time**
  - **Batch Mode – process a corpus of data assets**
- **Scope does not include storage, indexing or search functions over metacard contents.**

Army Participation

- **Active participation in the AMPS Working Group, the Information Assurance Subgroup, and the Spiral One Pilot Testing and Analysis. This participation included:**
  - **Contributing to the development of both general AMPS requirements and the information assurance requirements**
  - **Providing data assets based on the Blue Force Tracking (BFT) COI and the Battle Command Knowledge System (BCKS) for the test and evaluation activities.**
  - **Qualitative evaluation and feedback of the DDMS metacards created by the execution of the AMPS application**
  - **Feedback to and coordination with the AMPS technical team concerning installation and experimentation using the AMPS web service on a laptop**

Spiral 1 Scope – General

- **AMPS Working Group**
  - **Meeting/telecon biweekly at Arlington, VA between March and October 2007**
  - **Developed definitions, requirements, and scope for the service**
  - **Result was a thorough requirements specification [AMPS Working Group. AMPS Requirements v3, (18 October 2007)]**
- **Defined Scope:**
  - **Produce Discovery Metadata from COI Assets (Defense Readiness Service (DRS), Blue Force Tracking (BFT), Intelligence Agency (IA), Generic)**
  - **Exploit Open Standards**
  - **Label Metacards with security markings**
  - **Cryptographically Bind Metacards with Original**

Spiral 1 Scope - Corpus

- **Corpus by format and asset type**

	BFT	DRS	IA	Generic	Total
MS Word	33		40		**73**
HTML	65				**65**
TXT	5			4	**9**
OWL	1	5			**6**
PDF	4			342	**346**
MS PowerPoint	19	1			**20**
WSDL		6			**6**
MS Excel	37				**37**
XML	12				**12**
XSD	2	5			**7**
Total					**581**
Message Format	5			4	**9**
Email	57				**57**
PLI Rollup	30				**30**

Spiral 1 - Vocabularies

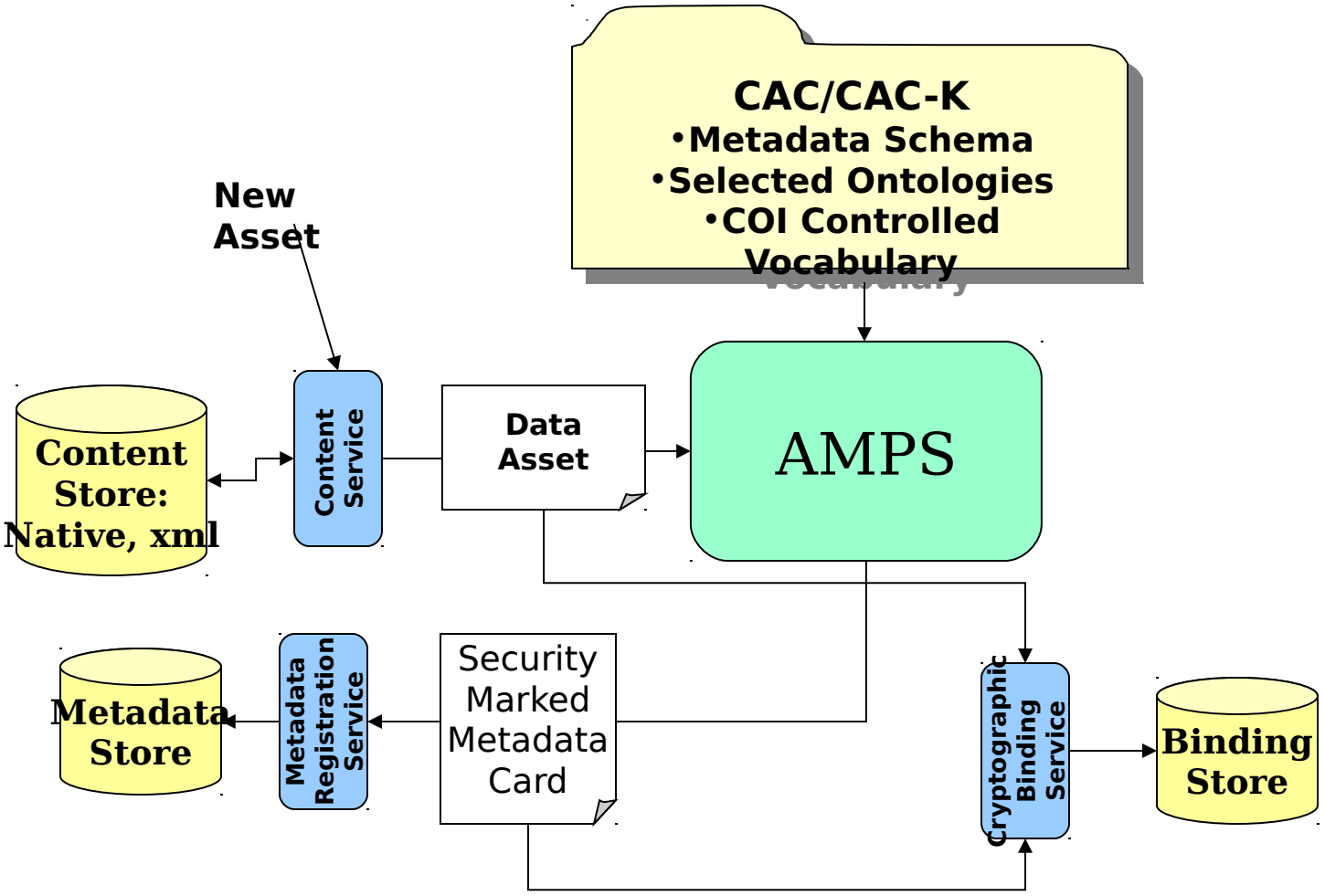
	BFT	DRS	Generic	Total
Entries	16	53	16655	**16724**
Synonyms	13	236	18875	**19124**
Assets	179	17	345	**541**

- **Volume does not equal quality/relevance**
- **Generic vocabulary from Defense Technical Information Center (DTIC) thesaurus**
  - **Broadly applicable to all Defense COIs**
  - **Ability to test scalability of vocabulary exploitation**
- **BFT & DRS very specific to COI information exchanges**

Spiral 1 Scope – DDMS Elements

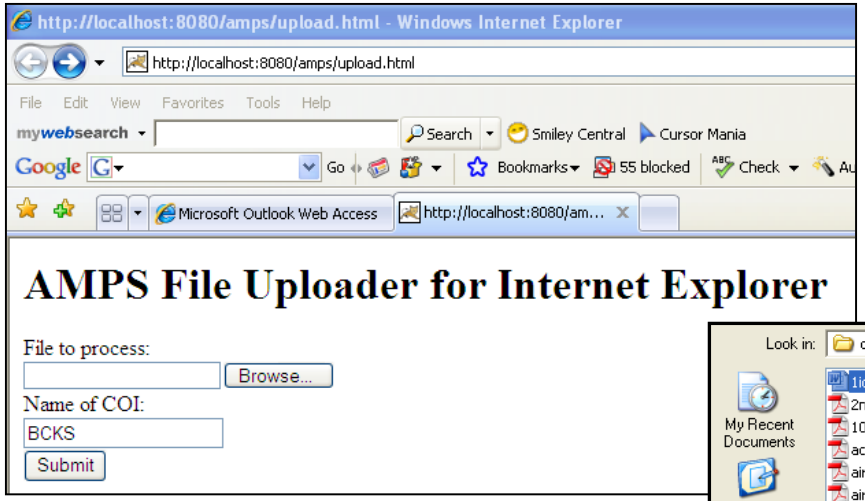
- Creator (mandatory , security classification required)
- Title (mandatory, security classification required)
- Subject (mandatory)
- Identifier (mandatory)
- Security (mandatory)
- Geospatial Coverage (mandatory unless not applicable)
- Date
- Format
- Type
- Description (security classification required)

AMPS Operational Example

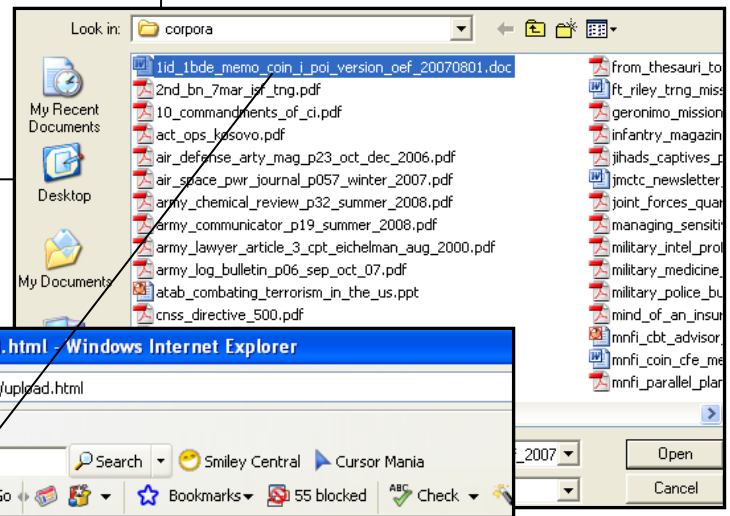


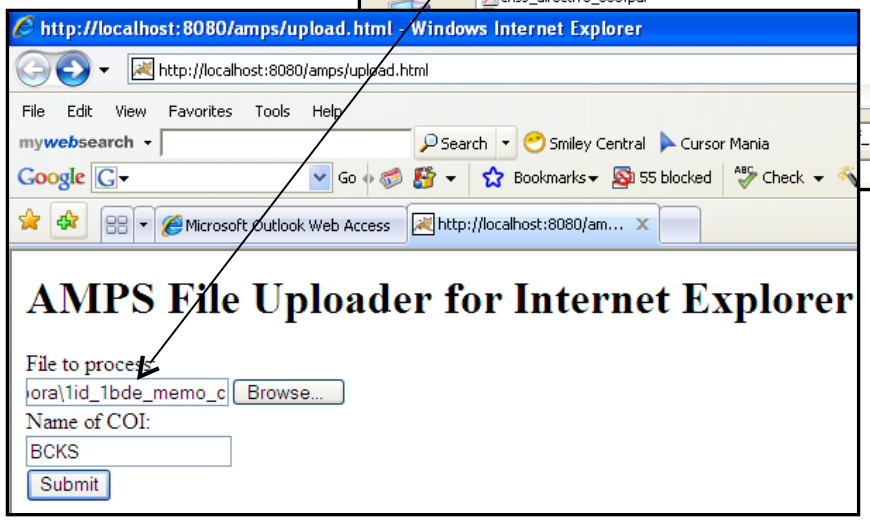
Single File AMPS Workflow

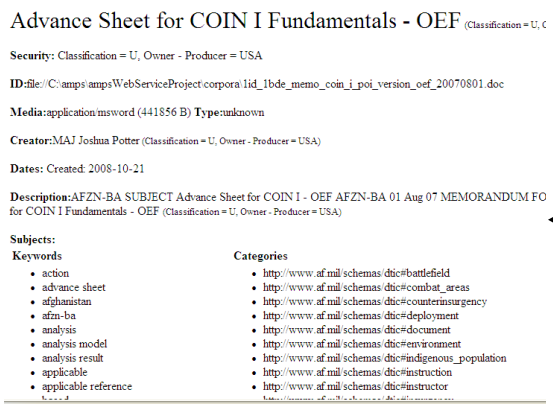
Open Apache Tomcat Server



Opens IE and the AMPS User Interface (UI)






Advance Sheet for COIN I Fundamentals - OEF (Classification = U, C

Security: Classification = U, Owner - Producer = USA

ID:file://C:\amps\ampsWebServiceProject\corpora\1id_1bde_memo_coin_i_poi_version_oef_20070801.doc

Media:application/msword (441856 B) **Type:**unknown

Creator:MAJ Joshua Potter (Classification = U, Owner - Producer = USA)

Dates: Created: 2008-10-21

Description:AFZN-BA SUBJECT Advance Sheet for COIN I - OEF AFZN-BA 01 Aug 07 MEMORANDUM FO for COIN I Fundamentals - OEF (Classification = U, Owner - Producer = USA)

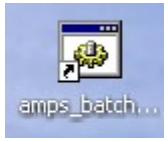
Subjects:

Keywords
- action
- advance sheet
- afghanistan
- afzn-ba
- analysis
- analysis model
- analysis result
- applicable
- applicable reference

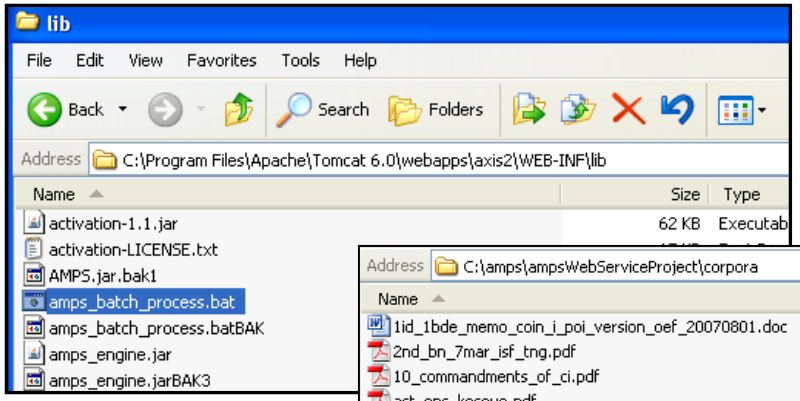
Categories
- http://www.af.mil/schemas/dtic#battlefield
- http://www.af.mil/schemas/dtic#combat_areas
- http://www.af.mil/schemas/dtic#counterinsurgency
- http://www.af.mil/schemas/dtic#deployment
- http://www.af.mil/schemas/dtic#document
- http://www.af.mil/schemas/dtic#environment
- http://www.af.mil/schemas/dtic#indigenous_population
- http://www.af.mil/schemas/dtic#instruction
- http://www.af.mil/schemas/dtic#instructor

Metacard Result

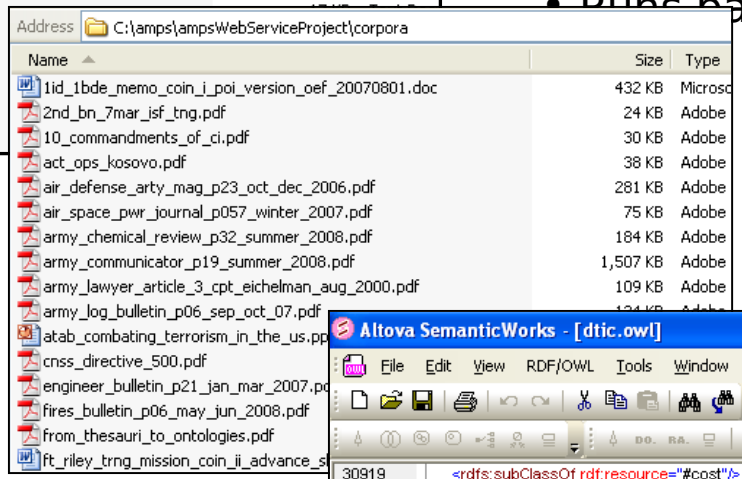
Batch Process AMPS Workflow

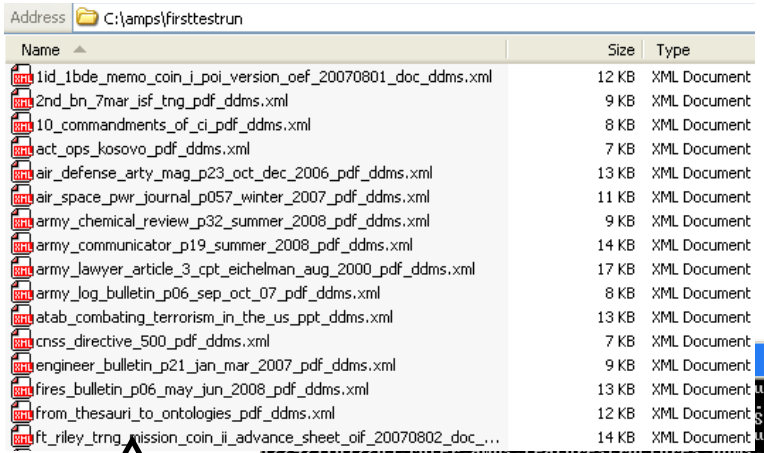


AMPS Batch Server

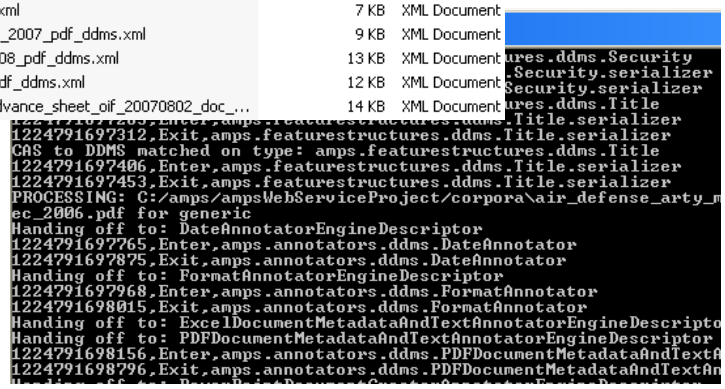


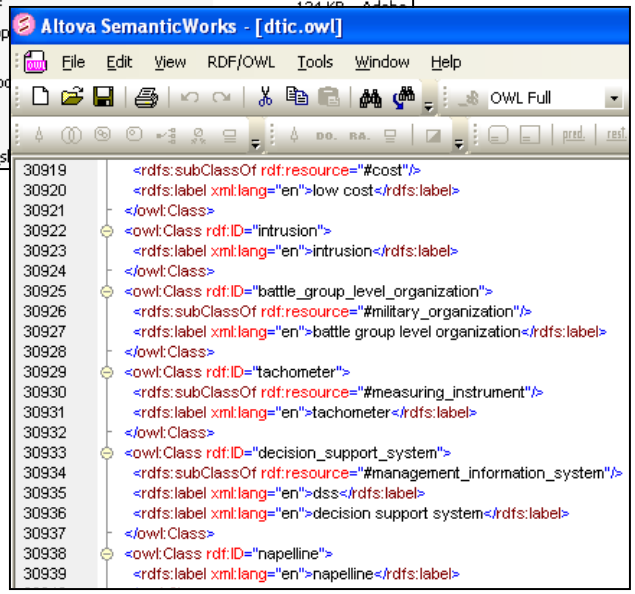
- Initiates AMPS
- Fetches files
- Applies an ontology
- Runs batch




Address C:\amps\firsttestrun
Name Size Type
1id_1bde_memo_coin_i_poi_version_oef_20070801_doc_ddms.xml 12 KB XML Document
2nd_bn_7mar_isf_tng_pdf_ddms.xml 9 KB XML Document
10_commandments_of_ci_pdf_ddms.xml 8 KB XML Document
act_ops_kosovo_pdf_ddms.xml 7 KB XML Document
air_defense_arty_mag_p23_oct_dec_2006_pdf_ddms.xml 13 KB XML Document
air_space_pwr_journal_p057_winter_2007_pdf_ddms.xml 11 KB XML Document
army_chemical_review_p32_summer_2008_pdf_ddms.xml 9 KB XML Document
army_communicator_p19_summer_2008_pdf_ddms.xml 14 KB XML Document
army_lawyer_article_3_cpt_eichelman_aug_2000_pdf_ddms.xml 17 KB XML Document
army_log_bulletin_p06_sep_oct_07_pdf_ddms.xml 8 KB XML Document
atab_combating_terrorism_in_the_us_ppt_ddms.xml 13 KB XML Document
cnss_directive_500_pdf_ddms.xml 7 KB XML Document
engineer_bulletin_p21_jan_mar_2007_pdf_ddms.xml 9 KB XML Document
fires_bulletin_p06_may_jun_2008_pdf_ddms.xml 13 KB XML Document
from_thesauri_to_ontologies_pdf_ddms.xml 12 KB XML Document
ft_riley_trng_mission_coin_ii_advance_sheet_oif_20070802_doc_... 14 KB XML Document

Produces XML



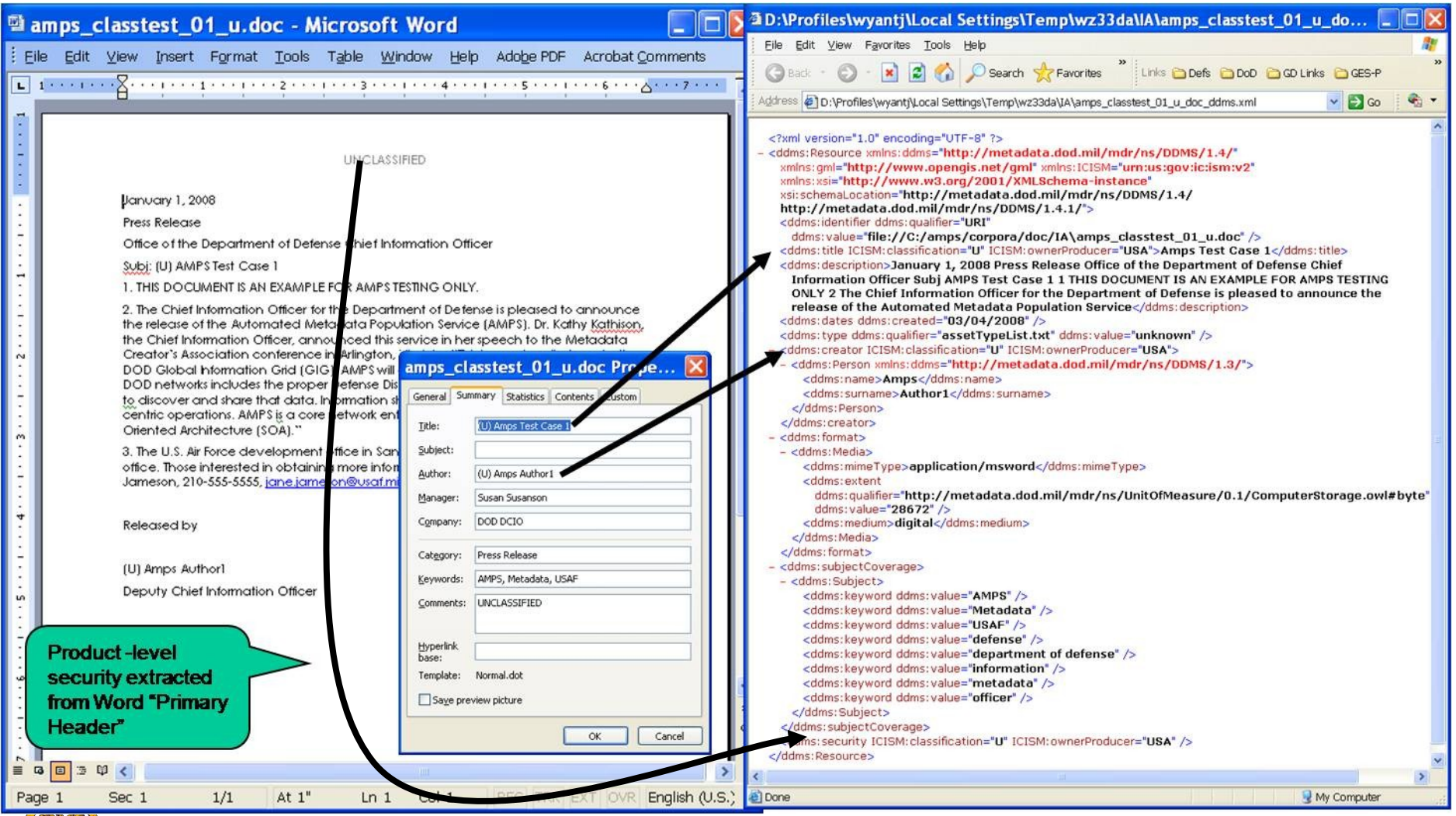


Metacard

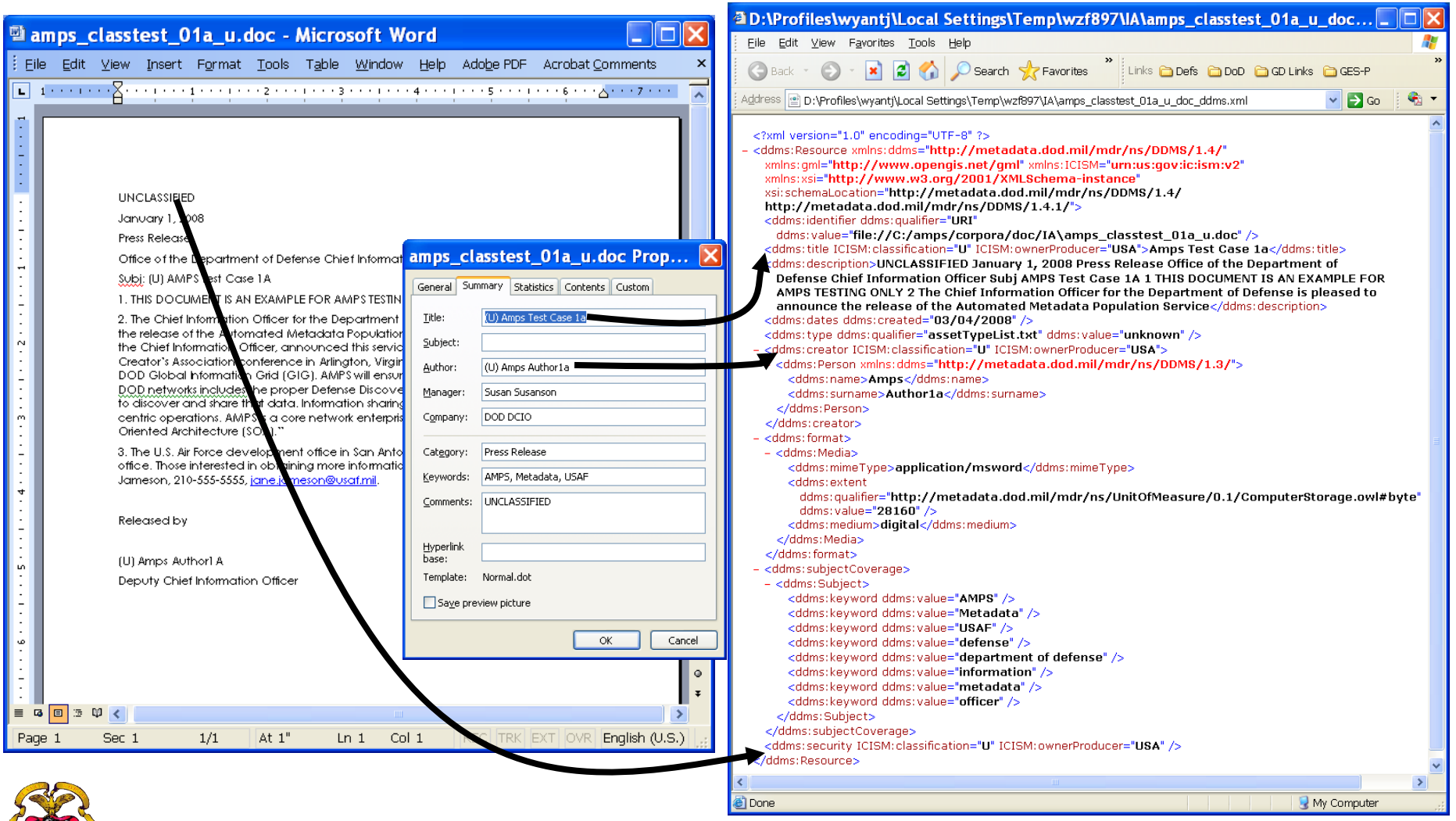
US Army Combined Arms Center

UNCLASSIFIED

Security Annotator – Sample 1



Security Annotator – Sample 2



Metacard Creation

DEPARTMENT OF THE ARMY
HSC 62nd ENGINNEER COMBAT BATTALION (HEAVY)
Fort Hood, Texas 76544

← Producer/Publisher

AFVG-EN-HSC

01September 2003

← Date Created

MEMORANDUM FOR HSC 62nd ENGINEER BATTALION

SUBJECT: Standing Operating Procedure for External Operations.

← Title

1. PURPOSE. To provide guidance and responsibilities for Direct Support maintenance procedures and policies. To standardize routine reoccurring maintenance operations, management procedures in performing its Direct Support mission.

2. SCOPE: This SOP is applicable to the 62nd Engineer Battalion.

3. OBJECTIVE. To set standards and goals in accordance with Army regulations and the Commanders intent.

   a. To provide guidelines for maintenance operations so that we as a support unit may provide effective and expeditious support to the customer.
   b. To develop an aggressive maintenance program that will ensure equipment is available for missions, operations and training requirements.
   c. To ensure operators, mechanics and supervisors are trained in their areas of maintenance responsibilities.
   d. To standardize maintenance management, and maintenance operations programs.
   e. To attain a satisfactory rating on all external maintenance evaluations
   f. To ensure all AOAP procedures are following in accordance with the regulations.

4. These procedures, which most cases already reflect current operation, will enhance the quality of Customer support provided by the 62nd EN BN Direct Support Activity.

5. CONCEPT. This SOP will describe DS level maintenance procedures and policies. It will also allow the commanders and maintenance managers of the unit the flexibility they need in designing and implementing an effective maintenance operations plan that will ensure maxinum operational readiness.

6. The point of contact is the Shop Officer at 288-7473

Keywords extracted from body of document

JOHN SENSLEY
CPT, EN
Commanding

← Creator/Author

BCKS Content Upload and Metadata Extraction

```
<?xml version="1.0" encoding="UTF-8" ?>
- <ddms:Resource xmlns:ddms="http://metadata.dod.mil/mdr/ns/DDMS/1.4/" xmlns:gml="http://www.opengis.net/gml"
  xmlns:ICISM="urn:us:gov:ic:ism:v2" xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xsi:schemaLocation="http://metadata.dod.mil/mdr/ns/DDMS/1.4/ http://metadata.dod.mil/mdr/ns/DDMS/1.4.1/">
  <ddms:identifier ddms:qualifier="URI"
    ddms:value="file://C:/amps/corpora\pdf\battleCommand\62_engr_bn_opns_sop_20030901.pdf" />
  <ddms:title ICISM:classification="U" ICISM:ownerProducer="USA">Standing Operating Procedure for External Operations.</ddms:title>
  <ddms:description ICISM:classification="U" ICISM:ownerProducer="USA">DEPARTMENT OF THE ARMY HSC 62 nd ENGINNEER COMBAT
    BATTALION Fort Hood, Texas 76544 AFVG-EN-HSC 01September 2003 MEMORANDUM FOR HSC 62nd ENGINEER BATTALION
    SUBJECT Standing Operating Procedure for External Operations</ddms:description>
  <ddms:dates ddms:created="2008-01-31" />
  <ddms:type ddms:qualifier="assetTypeList.txt" ddms:value="unknown" />
- <ddms:creator ICISM:classification="U" ICISM:ownerProducer="USA">
  - <ddms:Person>
      <ddms:name>JOHN</ddms:name>
      <ddms:surname>SENSLEY</ddms:surname>
    </ddms:Person>
  </ddms:creator>
- <ddms:format>
  - <ddms:Media>
      <ddms:mimeType>application/pdf</ddms:mimeType>
      <ddms:extent ddms:qualifier="http://metadata.dod.mil/mdr/ns/UnitOfMeasure/0.1/ComputerStorage.owl#byte"
        ddms:value="103962" />
      <ddms:medium>digital</ddms:medium>
    </ddms:Media>
  </ddms:format>
- <ddms:subjectCoverage>
  - <ddms:Subject>
      <ddms:keyword ddms:value="accepted" />
      <ddms:keyword ddms:value="accountability" />
      <ddms:keyword ddms:value="actions" />
      <ddms:keyword ddms:value="administration" />
      <ddms:keyword ddms:value="air supply" />
      <ddms:keyword ddms:value="alternator" />
      <ddms:keyword ddms:value="ammunition" />
      <ddms:keyword ddms:value="applicable" />
      <ddms:keyword ddms:value="appropriate" />
      <ddms:keyword ddms:value="ar 750" />
      <ddms:keyword ddms:value="ar 750-1" />
      <ddms:keyword ddms:value="automatic transmission" />
```

Title

Date Created

Producer/Publisher

Creator/Author

Keywords extracted from body of document

Keyword Metadata

- **What are the queries a searcher would use to get to this content?**

DEPARTMENT OF THE NAVY
CHIEF INFORMATION OFFICER
1000 NAVY PENTAGON
WASHINGTON, DC D3SD-1DDQ
20 October 2005
MEMORANDUM FOR DISTRIBUTION
Subj: DEPARTMENT OF THE NAVY KNOWLEDGE MANAGEMENT STRATEGY
1. Purpose. To communicate the Knowledge Management (KM) strategy for the Department of the Navy (DON) and identify over-arching focus areas for successful KM implementation.
2. Background. Knowledge Management is defined as the integration of people and processes, enabled by technology, to facilitate the exchange of operationally relevant information and expertise to increase organizational performance.
Knowledge Management is the cornerstone of decision superiority, knowledge dominance and information superiority. KM is a discipline that guides information and knowledge processes with the ultimate goal of helping to deliver mission effects. It is at the heart of force transformation. Three transforming DON documents, the DON IM/IT Strategic Plan 2006-2007, the Naval Transformation Roadmap 2003, and FORCEnet, A Functional concept for the 21s' Century cite the importance of knowledge sharing, actionable knowledge, mutual understanding, and collaborative decision-making. DoD's tenets of Network-Centric Warfare call for improved information sharing, enhanced quality of information, shared situational awareness, and collaboration that will result in increased mission effectiveness.

Keyword Extraction

DEPARTMENT OF THE NAVY
CHIEF INFORMATION OFFICER
1000 NAVY PENTAGON
WASHINGTON DC D3SD-1DDQ
20 October 2005
MEMORANDUM DISTRIBUTION
Subj DEPARTMENT OF THE NAVY KNOWLEDGE MANAGEMENT STRATEGY
1 Purpose to communicate the Knowledge Management (KM) strategy for the Department of the Navy (DON) and identify [redacted] arching focus areas for successful KM implementation
2 Background. Knowledge Management is defined as the integration of people and processes enabled by technology to facilitate the exchange of operationally relevant information and expertise to increase organizational performance
Knowledge Management is the cornerstone of decision superiority knowledge dominance and information superiority KM is a discipline that guides information and knowledge processes [redacted] the ultimate goal of helping to deliver mission effects. [redacted] the heart of force transformation Three transforming DON documents the DON IM/IT Strategic Plan 2006-2007 the Naval Transformation Roadmap 2003 and FORCEnet A Functional concept for the 21s' Century cite the importance of knowledge sharing actionable knowledge mutual understanding and collaborative decision-making DoD's tenets of Network-Centric Warfare call for improved information sharing enhanced quality of information shared situational awareness and collaboration [redacted] result in increased mission effectiveness

Documentation

- **AMPS Spiral 1**
  - **Requirements document**
  - **Technical Report**
  - **Developer's Guide – how to increase functionality**
  - **User's Guide – how to install in a new environment**
- **UIMA**
  - **Excellent tutorial for installation and use**

Getting Stuff to ~~Market~~

AMPS Workshop ~~Review~~

- Background
- DoD Discovery Metadata Specification (DDMS) DDMS
- AMPS Implementation and Functionality
- Army Participation in Spiral 1 Testing
- AMPS Operational Example
- Demonstration
- Documentation